

RESEARCH PAPER RP1072

Part of Journal of Research of the National Bureau of Standards, Volume 20,
February 1938

QUANTITATIVE ANALYSIS, WITH RESPECT TO THE COMPONENT STRUCTURAL GROUPS, OF THE INFRA- RED (1 TO 2 μ) MOLAL ABSORPTIVE INDICES OF 55 HYDROCARBONS¹

By Frank W. Rose, Jr.²

ABSTRACT

Data from previous studies on the infrared absorption spectra between 1.2 and 1.8 μ (5400 to 8900 cm^{-1}) for 54 hydrocarbons have been analyzed mathematically. The total molal absorptive index, K , at any wave length, is shown to follow, within the limits necessary to define the number of groups, the formula

$$K = n_a \alpha + n_b \beta \quad (a)$$

for the normal paraffin hydrocarbons, and the formula

$$K = n_a \alpha' + n_b \beta + n_c \gamma + n_d \delta \quad (b)$$

for other classes of hydrocarbons, where n_a , n_b , n_c , n_d are the number of $-\text{CH}_3$, $>\text{CH}_2$, $>\text{CH}$, and $>\text{CH}$ (aromatic) groups and α (or α'), β , γ , δ , are the values respectively, of the absorptive index for the respective unit group. Equation b is demonstrated to hold for branched-chain paraffins, cycloparaffins (naphthenes) and aromatic hydrocarbons, including 10 of high molecular weight (C_{24} to C_{32}). The accuracy of the method as well as restrictions on its application are given. Possible extensions in order to distinguish between paraffins and cycloparaffins are indicated. Its use is demonstrated by an example of the identification of an unknown isomeric nonane.

CONTENTS

	Page
I. Introduction.....	130
II. Discussion of concepts.....	130
III. Calculation of unit absorption coefficients.....	134
IV. Results.....	144
V. Discussion.....	148
1. Normal paraffins.....	148
2. Isoparaffins.....	148
3. Naphthene and aromatic hydrocarbons.....	150
4. Hydrocarbons of high molecular weight.....	151
VI. Determination of the number and kind of structural groups in unknown hydrocarbons.....	152
1. General procedure.....	152
2. Example: Determination of the structure of an unknown isomeric nonane.....	153
3. Limitations of the method.....	156
VII. References.....	156

¹ Financial assistance has been received from the research fund of the American Petroleum Institute. This work is part of Project 6, The Separation, Identification, and Determination of the Constituents of Petroleum.

² Research Associate at the National Bureau of Standards, representing the American Petroleum Institute.

I. INTRODUCTION

In connection with the work of the American Petroleum Institute Research Project 6 on the separation and identification of hydrocarbons from petroleum, the property of infrared absorption between 1.2 and 1.8 μ (5400 to 8900 cm^{-1}) was first used to "fingerprint" the hydrocarbons isolated [1]³ and later, in a qualitative way, to aid in the identification [2]. For the hydrocarbons of low molecular weight, most of which have been synthesized and subjected to measurement, the identification can be definitely made by comparison with the synthetic hydrocarbons, using such physical properties as molecular weight, carbon-hydrogen content, boiling point, freezing point, density, and refractive index, as discussed by Washburn [3], and leaving the infrared-absorption spectra as a corroborating "fingerprint." In the case of hydrocarbons of higher molecular weight, however, relatively few have been synthesized, and it often becomes impossible to identify the hydrocarbons isolated from petroleum because of the absence of data on the corresponding synthetic hydrocarbons of known structure.

In this latter case, recourse must be had to methods of identification that will yield a knowledge of the class or type of the given unknown molecule and of the nature and number of its structural units. A number of effective methods for classifying unknown hydrocarbons and their mixtures according to the general or average type of molecule have been developed and used extensively in recent years. These methods include the use of the molecular volume [4], the viscosity index [4, 5, 6, 7, 8], the specific refraction [9], the specific dispersion [9, 10, 11, 12, 13, 14], the viscosity-gravity index [15], refractivity intercept [16], etc. But known properties which yield information concerning the actual structural units of hydrocarbon molecules are rare. It has been known for several years that the infrared absorption due to vibrations between carbon and hydrogen atoms in the region of the spectrum from 1.2 to 1.8 μ is different in both wave length and intensity for $-\text{CH}_3$, $>\text{CH}_2$, $>\text{CH}$, and $>\text{CH}$ (aromatic) groups [17, 2].⁴ It seemed desirable, therefore, to extend the method of Wulf and Liddel [36] so that this property might yield quantitative, as well as qualitative, information concerning the structural units of an unknown hydrocarbon molecule.

The present report gives the results of a quantitative mathematical analysis of the infrared-absorption curves for 54 hydrocarbons that have already been published by the Project [2, 18], with the purpose of yielding the numerical values for the number of structural groups in a given unknown hydrocarbon molecule whose molecular weight and carbon-hydrogen composition are known. The method proposed is neither absolute nor independent, but reliable results can be obtained by the use of this property in conjunction with a judicious consideration of certain other properties of the unknown hydrocarbon.

II. DISCUSSION OF CONCEPTS

Although W. Herschel [19], in 1800, demonstrated the existence of radiation having wave lengths longer than those of the visible spectrum, the infrared-absorption spectra of organic compounds were not

³ Figures in brackets here and throughout the text refer to the references at the end of the paper.

⁴ Hereafter an incomplete valence bond indicates the attachment to the remainder of the molecule through another C atom.

studied until 1881, when Abney and Festing [20] investigated a series of typical organic liquids in the near infrared (0.7 to 1.2 μ) by photographic means. These authors correctly stated (for this region) "that the foundation of all absorptions in these bodies is the hydrogen," and furthermore advanced the hypothesis that absorption at specific wave lengths can be attributed to certain radicals. These authors also made a distinction between paraffinic and aromatic bonds. Using a bolometer, Julius [21] extended the range of measurements to 10 μ , and concluded that the absorption spectrum is a function not only of the elementary composition but also of the structure of the compound. Julius assigned the maxima at 3.45 μ and 6 to 7 μ ⁵ in the curves of percentage absorption plotted against wave length to the presence of methyl ($-\text{CH}_3$) groups in the molecule. Later Donath [23], from work on essential oils, found maxima at 1.69 and 2.2 μ and concluded that the action is intermolecular and not intramolecular, because the absorption bands were not as sharp as those in the visible spectra. Coblenz [22] stated that Ransohoff [24] measured the absorption spectra of six alcohols, finding maxima at 7.71, 3.0, and 3.43 μ , the absorption at 3.0 μ being assumed to be characteristic of the hydroxyl group. Puccianti [25] investigated a series of benzene derivatives, and found that all compounds which had carbon bound to hydrogen showed a maximum at 1.71 μ , while all molecules containing a benzene nucleus had additional maxima at 2.18 and 2.49 μ . From this evidence Puccianti concluded that the groups in the molecule are the governing influence in infrared absorption in this region. In 1905, Coblenz [22] published an extensive work on 135 compounds (solids, liquids, and gases, mainly organic) with the principal conclusions that:

1. The arrangement (bonding) of the atoms has a great influence on the resulting absorption spectrum.
2. No shifting of the maximum of absorption is observed with increase in molecular weight (except for gases).
3. A distinction must be made between a true shifting of the maximum and a shift which, on the introduction of a new adjacent band, occurs because the center of gravity is changed.
4. Certain absorbing groups preserve their identity in compounds containing other groups.

Later workers tended to confirm these results by Coblenz.⁶ In recent papers Barnes [27, 28], Bloch and Errera [44], and others have presented the modern conceptions of the relation of absorbing groups to infrared spectra, while Mecke [45, 46] has reviewed the use of the photographic infrared in the determination of constitutional formulas. Bartholomé and Teller [47] have made theoretical calculations relating the infrared spectra to the constitution of organic compounds.

The region, 6 to 15 μ , in which the infrared absorption is determined mainly by the carbon skeleton, has been the subject of much investigation by Lecomte and coworkers [26]. Thus Lambert and Lecomte [29] were able to make certain deductions concerning the nature of petroleum distillates by measurements in this region.

⁵ As corrected by Coblenz [22].

⁶ An excellent review of the development of this field is found in the recent treatise by Lecomte [26].

In 1919, Henri [30]⁷ attempted to correlate the intensity of absorption with molecular structure by assuming that the maximum at $3.4\ \mu$ in the absorption curves of alcohols was a linear function of the number of C—H bonds. Bonino [31], on the other hand, expressed the molal absorptive index as various functions (logarithmic, parabolic, and linear) of the number of such bonds, depending on the class of compound. Bonino's conclusions were in discord with those of Bell [32], who found that di-*n*-butyl and diisobutyl carbonates had different absorptive power although containing the same number of C—H bonds. However, the justification for this distinction became apparent when Brackett [17], using a high dispersion, resolved into its components the band occurring at $1.2\ \mu$ in the absorption spectra of hydrocarbons. Thus Brackett's work showed definitely that a distinction must be made between the $-\text{CH}_3$, $>\text{CH}_2$, and $>\text{CH}$ groups. Later, Liddel and Kasper [2] not only confirmed these findings but also showed that "aromatic" absorption could be distinguished. In the same report of Liddel and Kasper, similar results were obtained for the bands at 1.8 and $1.4\ \mu$. Freymann [33] found analogous behavior at 0.9 and $1\ \mu$. A previous paper [18] by the author extended the scope of the compounds measured to hydrocarbons of high molecular weight. It is these units, $-\text{CH}_3$, $>\text{CH}_2$, $>\text{CH}$, and $>\text{CH}$ (aromatic), which are the variables in terms of which the intensity of absorption must be expressed.

Meanwhile, Liddel and Wulf [34] stated that "First, the absorption coefficient for the N—H bond in its absorption at $1.5\ \mu$ throughout a very considerable range in the character of the compounds . . . remains the same order of magnitude. Second, there are measurable differences in this absorption coefficient from class to class; i. e., primary aliphatic, secondary aliphatic, primary aromatic, etc., as well as differences in the position of the maximum of absorption. Third, when a normal paraffin greater than two carbon atoms long is attached to the amine group,⁸ the absorption coefficient and maximum of absorption vary very little." In a later paper, Wulf and Liddel [36] use the area under the molal absorptive index curve in a given region as a quantitative measure of the number of bonds which give rise to the absorption at that place.

If, for various compounds, the form of the absorption curve is identical, the ordinate of the curve of the molal absorptive index at a given wave length is directly proportional to the total area under the curve for a given region [36]. It follows, under these conditions, that the ordinate at a given wave length, preferably near the maximum of a given "hump", will be proportional to the number of groups giving rise to the absorption at that place.⁹ To test the Wulf and Liddel assumption that the area is proportional to the number of groups, and also to check, in part, the assumed similarity in the

⁷ The insistence of Henri upon the use of the molal absorption index instead of the percentage absorption or transmission led to more readily apparent relationships with molecular structure. The molal absorption index, K , is defined [48] thus:

$$K = \frac{-\log_{10} I/I_0}{cd},$$

where I is the intensity of the transmitted light, I_0 the incident light, c the concentration in moles per liter, and d the cell depth in centimeters.

⁸ Confirmation of this fact is given by the work of Rossini [35] on the heats of combustion of normal paraffins and alcohols where it is shown that intramolecular action is minor beyond two carbon atoms.

⁹ For the present purpose it is unnecessary to resolve the curve of an ideal group into components which can be associated with different modes of vibration, as Wulf [37] has done. It suffices to develop a value of the molal absorptive index for an average unit group at a fixed wave length.

form of the curve, a plot was made from the data on the normal paraffin hydrocarbons, of the value of the ordinate of the molal absorptive index curve at 8160, 8254, 8400, and 8750 cm^{-1} , against the number of $>\text{CH}_2$ groups in the molecule.¹⁰

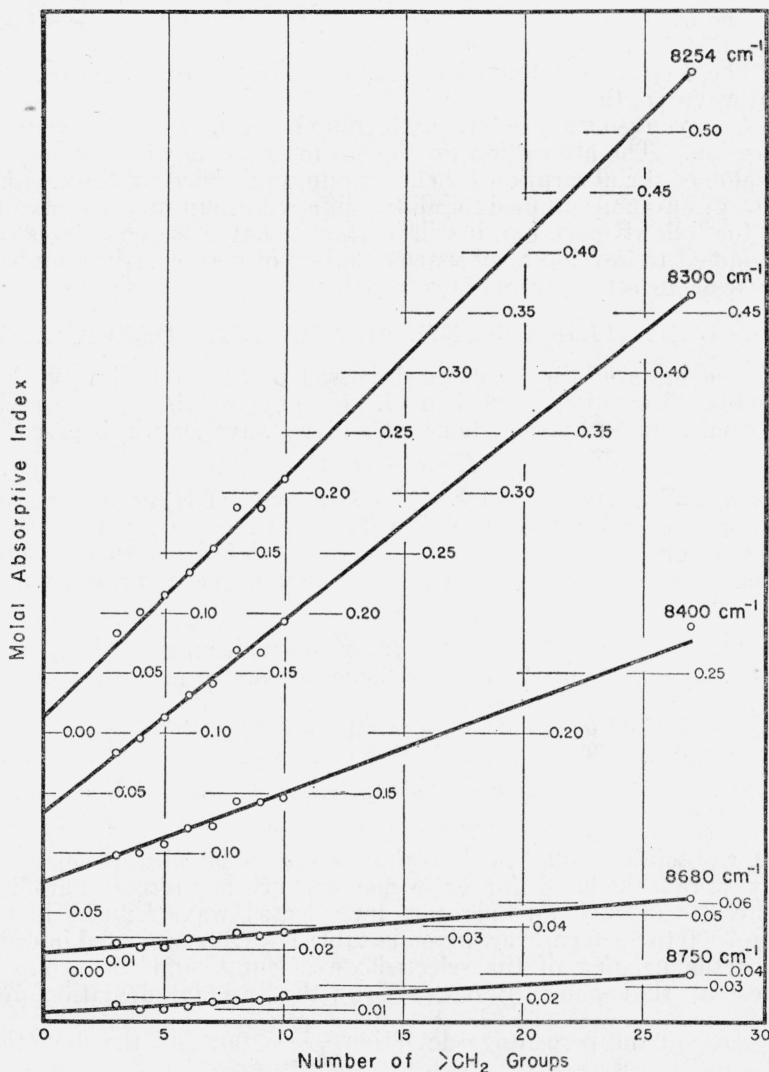


FIGURE 1.—Plot of the data on the absorption of the normal paraffin hydrocarbons at various wave numbers.

The scale of ordinates gives K , the molal absorption index, of the given normal paraffin at the wave length indicated at the right-hand side of the plot. The scale of abscissas gives n , the number of methylene ($>\text{CH}_2$) groups in the molecule.

Figure 1 shows that the given ordinate at the respective wave numbers is substantially linear with the number of $>\text{CH}_2$ groups

¹⁰ In the previous paper [18], a table is given showing the wave number at which maxima occur for the various structural groups.

in the molecule.¹¹ The positive intercept at zero number of $>\text{CH}_2$ groups represents the substantially constant absorption due to the two $-\text{CH}_3$ groups in each of the molecules.

The concepts, which have been discussed above and which form the initial basis of the present analysis, may be summarized as follows:

1. The hydrocarbons studied consist of the structural units $-\text{CH}_3$, $>\text{CH}_2$, $>\text{CH}$, and $>\text{CH}$ (aromatic).

2. These structural units have their maximum absorption at different wave lengths.

3. At any given wave length, each group has a constant unit value of absorption. The absorption for the entire molecule will therefore be the value of the absorption for each group, multiplied by the number of such groups and summed for all the different groups in the molecule.

In the following section, it will be shown that these postulates can be modified to develop a consistent method of mathematical analysis of the near-infrared spectra of hydrocarbons.

III. CALCULATION OF UNIT ABSORPTION COEFFICIENTS

On the basis of the concepts discussed in the preceding section, the molal absorption coefficient, K , for any member of the series of normal paraffin hydrocarbons, at a given wave length, is given by

$$K = n_a \alpha + n_b \beta, \quad (1)$$

where n_a and n_b are the number of $-\text{CH}_3$ and $>\text{CH}_2$ groups, respectively, and α and β are the unit values of the absorption for these respective groups at a given wave length. For the normal paraffin hydrocarbons, n_a equals 2, and eq 1 may properly be written as

$$K = 2\alpha + n_b \beta \quad (1a)$$

With known values of K and n_b , for given molecules at given wave lengths, the values of the unit coefficients α and β may then be calculated.

The method of least squares was selected for evaluating the coefficients α and β .¹² This involved solving the equations

$$\Sigma n_a K - \alpha \Sigma n_a^2 - \beta \Sigma n_a n_b = 0 \quad (2)$$

$$\Sigma n_b K - \alpha \Sigma n_a n_b - \beta \Sigma n_b^2 = 0 \quad (3)$$

After obtaining α and β , the values of the molal absorption coefficients were calculated for each member of the normal paraffins. The results of such an operation for selected wave lengths in the region 7900 to 8900 cm^{-1} are given in table 1. Here are listed in order (under the heading of the selected wave length and the group of interest at that point¹³): $K_{\text{obs.}}$; $K_{\text{calc.}}$; the residual deviation, $\Delta =$

$K_{\text{obs.}} - K_{\text{calc.}}$; the percentage deviation $\frac{\Delta}{K_{\text{obs.}}} 100$; and the deviation

in terms of the number of the group sought at that point.¹⁴

¹¹ After this work was completed, a paper by Williams [38] appeared in which the data of Coblentz [22] on the transmission at 3.4 μ for a series of fractions isolated from petroleum were converted to molal absorption coefficients and plotted as a function of the number of C-H bonds.

¹² The simpler method of averages (Lipka [39], p. 126) was not used since it involves division of the data into two groups. Experience has shown that different groupings often lead to inconsistent results. Thus complications might result if it be true, as suggested by Liddel and Kasper [2], that maxima at certain wave lengths show an alternation in wave length or position with the number of carbon atoms in the chain.

¹³ The wave numbers at which data were taken were so selected that some one of the various types of structural groups possessed a maximum of absorption at or near such wave numbers. As a result, the numerical value for the group of interest could be calculated with accuracy.

¹⁴ The derivations of unit coefficients γ and δ for $>\text{CH}$ and $>\text{CH}$ (aromatic) are explained below. Deviations in terms of such groups are included in table 1 to show that an analysis would not falsely indicate the presence of such groups.

TABLE 1.—Results of representing, for the normal paraffins, the total molar absorption coefficient, K , as a function of the number, n_a , of primary ($-\text{CH}_3$), and, n_b , of secondary ($>\text{CH}_2$) groups

$$K = n_a \alpha + n_b \beta$$

Compound	Number of groups, $>\text{CH}_2$	8160 cm^{-1} ($>\text{CH}$)					8254 cm^{-1} ($>\text{CH}_2$)					8400 cm^{-1} ($-\text{CH}_3$)				
		$\alpha=0.00295 \beta=0.00992 \gamma=0.01132$					$\alpha=0.00629 \beta=0.01989$					$\alpha=0.03770 \beta=0.00735$				
		n_b	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$	$\Delta = \text{Number}$ α of $>\text{CH}$ gr.	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$	$\Delta = \text{Number}$ β of $>\text{CH}_2$ gr.	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$
<i>n</i> -Pentane...	3	0.0469	0.0357	+0.0113	+31.7	+1.2	0.0822	0.0723	+0.0099	+12.1	+0.5	0.0978	0.0974	+0.0004	+0.4	0.0
<i>n</i> -Hexane...	4	.0443	.0456	-.0013	-2.9	-1	.0991	.0921	+.0070	+7.1	+3	.0992	.1048	-.0056	-5.7	-1
<i>n</i> -Heptane...	5	.0490	.0555	-.0065	-11.7	-7	.1147	.1120	+.0027	+2.4	+1	.1072	.1121	.0049	-4.6	-1
<i>n</i> -Octane...	6	.0667	.0654	+.0013	+2.0	+1	.1338	.1319	+.0019	+1.4	+1	.1217	.1195	+.0022	+1.8	+1
<i>n</i> -Nonane...	7	.0691	.0753	-.0062	-8.2	-7	.1534	.1518	+.0016	+1.1	+1	.1225	.1268	-.0043	-3.5	-1
<i>n</i> -Decane...	8	.0886	.0853	+.0033	+3.9	+4	.1877	.1717	+.0160	+8.5	+8	.1430	.1342	+.0088	+6.2	+2
<i>n</i> -Undecane...	9	.0985	.0952	+.0033	+3.5	+4	.1875	.1916	-.0041	-2.1	-2	.1430	.1415	-.0090	-6.3	-2
<i>n</i> -Dodecane...	10	.0995	.1051	-.0056	-5.3	-6	.2115	.2115	.0000	0.0	0	.1460	.1489	-.0029	-2.0	-1
<i>n</i> -Nonacosane	27	.2763	.2737	+.0026	+0.9	+3	.5500	.5495	+.0005	+1	0	.2893	.2739	+.0155	+5.4	+4
Average deviation					± 7.8	± 0.5				± 3.9	± 0.2				± 4.0	± 0.1

Compound	Number of groups, $>\text{CH}_2$	8300 cm^{-1} ($>\text{CH}_2$)					8680 cm^{-1} ($-\text{CH}_3$)					8750 cm^{-1} ($>\text{CH}$) (aromatic)				
		$\alpha=0.01681 \beta=0.01585$					$\alpha=0.00811 \beta=0.00176$					$\alpha=0.00327 \beta=0.00126 \delta=0.01121$				
		n_b	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$	$\Delta = \text{Number}$ β of $>\text{CH}_2$ gr.	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$	$\Delta = \text{Number}$ α of $-\text{CH}_3$ gr.	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}}-K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}}\cdot 100$
<i>n</i> -Pentane...	3	0.0836	0.0812	+0.0024	+2.9	+0.2	0.0241	0.0215	+0.0026	+10.8	+0.3	0.0138	0.0103	+0.0035	+25.4	+0.3
<i>n</i> -Hexane...	4	.0948	.0970	-.0022	-2.3	-1	.0218	.0233	-.0015	-6.9	-2	.0094	.0116	-.0022	-23.4	-2
<i>n</i> -Heptane...	5	.1122	.1128	-.0006	-0.5	0	.0204	.0251	-.0047	-23.0	-6	.0095	.0128	-.0033	-34.8	-3
<i>n</i> -Octane...	6	.1318	.1287	+.0031	+2.4	+2	.0287	.0263	+.0024	+6.6	+2	.0130	.0141	-.0010	-7.7	-1
<i>n</i> -Nonane...	7	.1405	.1445	-.0040	-2.8	-3	.0276	.0285	-.0009	-3.2	-1	.0147	.0154	-.0007	-4.7	-1
<i>n</i> -Decane...	8	.1691	.1604	+.0087	+5.4	+6	.0339	.0302	+.0037	+10.9	+5	.0167	.0166	+.0001	+ .6	0
<i>n</i> -Undecane...	9	.1670	.1762	-.0092	-5.5	-6	.0302	.0320	-.0018	-5.6	-2	.0175	.0179	-.0004	-2.3	0
<i>n</i> -Dodecane...	10	.1934	.1921	+.0013	+0.1	+1	.0348	.0338	+.0010	+2.9	+1	.0216	.0191	+.0025	+11.6	+2
<i>n</i> -Nonacosane	27	.4620	.4615	+.0005	+1	0	.0633	.0636	-.0003	-0.5	0	.0402	.0406	-.0004	-1.0	0
Average deviation					± 2.4	± 0.2				± 7.8	± 0.2				± 12.4	± 0.1

At 8254 cm^{-1} , only the values for *n*-octane to *n*-nonacosane were used (omitting that for *n*-decane, which is probably in error). This necessary modification of the basic concept 3, p. 134, was introduced because the residuals (excluding *n*-decane) manifest an increasing positive trend as the length of the chain is decreased.¹⁵ A possible explanation for this fact is given in the next section. Exclusion of the values of *K* for *n*-pentane, *n*-hexane, *n*-heptane, and *n*-decane altered α , while β remained relatively unchanged. Thus, the entire series yielded $\alpha=0.0095$, $\beta=0.0196$, while the selected points gave $\alpha=0.0063$, $\beta=0.0199$. It is important to note that, though the residual for *n*-pentane at 8254 cm^{-1} is high, the error in terms of the number of the groups sought is not large.

At 8750 cm^{-1} , the large deviations of *n*-pentane, *n*-hexane, and *n*-heptane caused their exclusion, though the values of the coefficients were not greatly altered by the omission. The use of the entire series yielded $\alpha=0.00307$, $\beta=0.00128$, while the use of the higher members gave $\alpha=0.00327$, $\beta=0.00126$.

With the exception of the values at 8254 and 8750 cm^{-1} , the values of the coefficients, α and β , were derived from the entire series of normal paraffins. Since the deviations of the lower members at all other wave numbers are for the most part random, this portion of the data confirms the simple concepts 1 to 3, as stated in section II.

The other regions, 5400 to 6400 and 6400 to 7400 cm^{-1} were similarly treated using all data, and yielded results similar to those in table 1. The values of α derived for these regions are as follows: 5668 cm^{-1} , 0.1056; 5781 cm^{-1} , 0.1078; 5800 cm^{-1} , 0.1624; 5860 cm^{-1} , 0.1970; 5870 cm^{-1} , 0.1914; 5900 cm^{-1} , 0.1574; 6950 cm^{-1} , 0.00909; 6975 cm^{-1} , 0.01133; 7064 cm^{-1} , 0.01253; 7185 cm^{-1} , 0.02138; and 7350 cm^{-1} , 0.01025. The corresponding values of β are those given in table 5. For consistency in calculation, the values of the unit coefficients given throughout the paper are expressed in more significant figures than are warranted by the data. An optimistic estimate of the inaccuracy of these coefficients would be 3 to 4 percent, while the conservative value used here in analysis for the experimental error of the total molar absorption coefficient is ± 8 percent.

The coefficients, α and β , were next applied to the branched-chain paraffins, as shown in the first portion of table 2 for the wave numbers 8400 and 5860 cm^{-1} . Consistently negative deviations increasing in magnitude with the increase in number of methyl groups indicate that the value for α is too large. On inspection, no simple variation in these residuals as a function of the number of secondary ($>\text{CH}_2$) groups was apparent. Hence although the basic concept 3, p. 134 (section II) does not seem usefully valid for the methyl groups, the simplifying assumption that the value of β remains unchanged for the isoparaffins was retained.¹⁶ It is shown later that a value, α' , modified to fit the data on the isoparaffins is applicable to all the classes of hydrocarbons considered, with few exceptions. Since the identity of

¹⁵ An equation containing three constants,

$$K = \alpha' n_a + \beta' n_b + \frac{\gamma'}{n_b} \quad (4)$$

was also tested for representation of the data on normal paraffins at 8254 cm^{-1} . Here the last term might be expected to eliminate the deviation for small values of n_b . But the residuals were practically unchanged.

¹⁶ That even this is not strictly true is demonstrated when a least-squares solution is applied to the isoparaffins alone. At 8400 cm^{-1} , the coefficients thus developed were $\alpha'''=0.0260$, $\beta'''=0.0106$. In this series, the value for the unit methyl group is probably more accurately evaluated than was the case for the normal series.

the normal paraffins need never be in doubt, the efficacy of this method for the determination of the structure of hydrocarbons is not impaired and the derivation of α may be considered as merely a step in the process of obtaining the more general coefficient, α' .

It was possible, at certain wave lengths, to use α unchanged, since the residuals manifested no trends. In the event that α , from the normal paraffins, seemed unsuitable for the calculations of the total molal-absorption index, the method adopted (except at 8160 and 6950 cm^{-1}) for deriving a new unit coefficient, α' , for a methyl group in the isoparaffins was as follows: After the effect of the secondary ($> \text{CH}_2$) linkages (as determined from the normal paraffins) was deducted, the sum of the remainders was divided by the total number of methyl groups, to yield a mean value for a single methyl group, α' , thus

$$\alpha' = \frac{\Sigma(K_{(obs.)} - n_b)}{\Sigma n_a} \quad (5)$$

This simple method yielded residuals for the isoparaffins that were not significantly larger than those obtained by the least-squares solution (see footnote 16, p. 136). The result of applying α' and β to the isoparaffins at 8400 and 5860 cm^{-1} is found in the second or revised portion of table 2.

TABLE 2.—Results of representing, for the branched-chain paraffins, the total molal absorption coefficient, K , as a function of the number, n_a , of primary ($-\text{CH}_3$), and, n_b , of secondary groups

$$K = n_a \alpha + n_b \beta$$

Compound	Number of groups		8400 cm^{-1}						5960 cm^{-1}				
	$-\text{CH}_3$	$>\text{CH}_2$	$\alpha=0.03770$			$\beta=0.00735$			$\alpha=0.01970$			$\beta=0.0194$	
	n_a	n_b	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \cdot 100$	$\frac{\Delta}{\alpha} = \text{Number of } -\text{CH}_3 \text{ gr.}$		$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \cdot 100$	$\frac{\Delta}{\alpha} = \text{Number of } -\text{CH}_3 \text{ gr.}$
2-Methylbutane.....	3	1	0.0891	0.1204	-0.0313	-35.1	-0.8		0.396	0.610	-0.214	-54.3	-1.1
2-Methylpentane.....	3	2	.1180	.1278	-.0098	-8.3	-.3		.524	.630	-.106	-20.3	-.6
2-Methylhexane.....	3	3	.1150	.1351	-.0201	-17.5	-.5		.494	.649	-.155	-31.4	-.8
2-Methylnonane.....	3	6	.1341	.1572	-.0231	-17.2	-.6		.550	.707	-.157	-28.6	-.8
3-Methylpentane.....	3	2	.1058	.1278	-.0220	-20.8	-.6		.490	.630	-.140	-28.6	-.7
3-Methylhexane.....	3	3	.1140	.1351	-.0211	-18.5	-.6		.512	.649	-.137	-26.8	-.7
3-Methylnonane.....	3	6	.1346	.1572	-.0226	-16.8	-.6		.580	.707	-.127	-21.9	-.6
4-Methylnonane.....	3	6	.1362	.1572	-.0210	-15.4	-.6		.557	.707	-.150	-26.9	-.8
5-Methylnonane.....	3	6	.1177	.1572	-.0395	-33.6	-1.0		.578	.707	-.129	-22.3	-.7
3-Ethylpentane.....	3	3	.1219	.1351	-.0132	-10.8	-.4		.513	.649	-.136	-26.5	-.7
3,3-Dimethylpentane.....	4	2	.1227	.1655	-.0431	-36.0	-1.1		.545	.827	-.282	-51.7	-1.4
2,3-Dimethylbutane.....	4	0	.1108	.1508	-.0400	-36.1	-1.0		.517	.789	-.272	-52.6	-1.4
2,3-Dimethylpentane.....	4	1	.1162	.1581	-.0420	-36.1	-1.1		.551	.807	-.256	-46.5	-1.3
2,4-Dimethylpentane.....	4	1	.1100	.1581	-.0481	-43.7	-1.3		.602	.807	-.205	-34.1	-1.0
2,2,3-Trimethylbutane.....	5	0	.1195	.1885	-.0653	-54.7	-1.7		.446	.985	-.539	-120.6	-2.7
2,2,4-Trimethylpentane.....	5	1	.1346	.1958	-.0612	-45.5	-1.6		.611	1.004	-.393	-64.3	-2.0
3,5,5-Trimethylheptane.....	5	3	.1522	.2106	-.0584	-38.3	-1.5		.752	1.043	-.291	-38.7	-1.5
3,4,5,5-Tetramethylhexane.....	6	1	.1502	.2335	-.0833	-55.5	-2.3		.765	1.201	-.436	-57.0	-2.2
Average deviation.....						-30.0	-1.0					-41.8	-1.2

Compound	Number of groups		8400 cm ⁻¹					5890 cm ⁻¹				
	-CH ₃	>CH ₂	$\alpha' = 0.02705 \quad \beta = 0.00735$					$\alpha' = 0.1363 \quad \beta = 0.0194$				
	n_a	n_b	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \cdot 100$	$\frac{\Delta}{\alpha'} = \text{Number of } -\text{CH}_3 \text{ gr.}$	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \cdot 100$	$\frac{\Delta}{\alpha'} = \text{Number of } -\text{CH}_3 \text{ gr.}$
2-Methylbutane.....	3	1	0.0891	0.0886	+0.0005	+0.6	0.0	0.396	0.428	-0.032	-8.1	-0.2
2-Methylpentane.....	3	2	.1180	.0959	+.0221	+18.7	+.8	.524	.448	+.076	+14.5	+.6
2-Methylhexane.....	3	3	.1150	.1033	+.0117	+10.2	+.4	.494	.467	+.027	+4.9	+.2
2-Methylnonane.....	3	6	.1341	.1253	+.0088	+6.6	+.3	.550	.525	+.025	+4.5	+.2
3-Methylpentane.....	3	2	.1058	.0959	+.0099	+9.4	+.4	.490	.448	+.043	+8.8	+.3
3-Methylhexane.....	3	3	.1140	.1033	+.0107	+9.4	+.4	.512	.467	+.045	+8.8	+.3
3-Methylnonane.....	3	6	.1346	.1253	+.0093	+6.9	+.3	.580	.525	+.055	+9.5	+.4
4-Methylnonane.....	3	6	.1362	.1253	+.0109	+8.0	+.4	.557	.525	+.032	+5.7	+.2
5-Methylnonane.....	3	6	.1177	.1253	-.0077	-6.6	-.3	.578	.525	+.053	+9.2	+.4
3-Ethylpentane.....	3	3	.1219	.1033	+.0186	+15.3	+.7	.513	.467	+.046	+9.0	+.3
3,3-Dimethylpentane.....	4	2	.1227	.1229	-.0002	-.2	.0	.545	.583	-.038	-7.0	-.3
2,3-Dimethylbutane.....	4	0	.1108	.1086	+.0026	+2.3	+.1	.517	.545	-.028	-5.4	-.2
2,3-Dimethylpentane.....	4	1	.1162	.1156	+.0006	+.5	.0	.551	.564	-.013	-2.4	-.1
2,4-Dimethylpentane.....	4	1	.1100	.1156	-.0056	-5.1	-.2	.602	.564	+.038	+6.3	+.3
2,2,3-Trimethylbutane.....	5	0	.1195	.1353	-.0158	-13.2	-.6	.446	.682	-.204	(-68.2)	(-2.2)
2,2,4-Trimethylpentane.....	5	1	.1346	.1426	-.0080	-6.0	-.3	.611	.701	-.099	-16.2	-.7
3,5,5-Trimethylheptane.....	5	3	.1522	.1573	-.0051	-3.4	-.2	.752	.740	+.012	+1.6	+.1
3,4,5,5-Tetramethylhexane.....	6	1	.1502	.1696	-.0194	-12.9	-.7	.765	.837	-.072	-9.4	-.5
Average deviation.....						± 7.5	± 3				± 7.7	± 3

Consideration of the tertiary ($>\text{CH}$) groups present in the isoparaffins as absorbing entities extends the equation analogous to eq 1 to the form

$$K = n_a\alpha' + n_b\beta + n_c\gamma, \quad (6)$$

where K , n_a , n_b , α' , and β follow the previous notation, and n_c represents the number of tertiary groups having the unit coefficient of absorption, γ . The coefficient, γ , was evaluated only at 8160 cm^{-1} and 6950 cm^{-1} ; at all other wave numbers it was assumed to be zero.

The disregard of $>\text{CH}$ as an absorbing group is justified by the weak absorption of the group at most wave numbers. Furthermore, the residuals at all wave numbers, examined with regard to a systematic variation of the residuals as a function of the number of $>\text{CH}$ groups, showed no significant trends. Indeed, this group has only a weak absorption, even at its maxima (8160 and 6950 cm^{-1}), since it involves only one C—H bond, and hence, at other wave numbers, becomes very small.

At 8160 cm^{-1} , a preliminary value for the unit coefficient, γ , was obtained by subtracting from the observed molal absorption coefficient the absorption due to $-\text{CH}_3$ and $>\text{CH}_2$, using α and β , the unrevised unit coefficients derived from the normal paraffins. The remainders were, in all but a few cases, positive, and varied in magnitude with the number of $>\text{CH}$ groups in the molecule. The sum of these remainders divided by the sum of the tertiary $>\text{CH}$ groups yields the unit coefficient, γ .

Another value was obtained by subtracting from each observed molal absorption coefficient only the absorption attributable to secondary ($>\text{CH}_2$) groups (assuming β to be unchanged, as was done in the development of α') and applying a least-squares solution for α' and γ to the remainders. Thus, in the equation

$$K_{(\text{obs.})} - \beta n_b = n_a\alpha' + n_c\gamma, \quad (7)$$

only α' and γ are unknown. Better agreement between observed and calculated values was obtained by the latter method, although the constant, γ , was not radically different. In the former method, using $\alpha = 0.00295$ and $\beta = 0.00992$, γ became 0.00921; while in the latter, with β remaining unchanged, and $\alpha' = 0.00197$, γ became 0.01132.

Application of the first method to the data for 6950 cm^{-1} yielded only slightly negative residuals, which indicated that the change in the unit coefficient for methyl ($-\text{CH}_3$) was sufficient to cloak the effect of the tertiary group. However, the second method yielded reasonable unit coefficients (see table 5) and these were adopted.

Before developing δ , the unit coefficient for $>\text{CH}$ (aromatic) absorption, it is necessary to define clearly the group to which it refers. Since, in the region 5400 to 8900 cm^{-1} , the absorption is due solely to the vibration of the C—H bonds, a benzene derivative will not show the same strength attributable to "aromatic" bonds as does the unsubstituted benzene nucleus if the absorption per $>\text{CH}$ (aromatic) bond remains substantially constant. Furthermore, the magnitude of the absorption decreases with the number of nuclear substitutions. Hence δ is the unit coefficient for one $>\text{CH}$ (aromatic) bond, of which benzene possesses six, toluene five, xylene four, etc. In this report, the possible consequences of the symmetry of benzene in rendering the absorption spectra unique are disregarded.

The numerical value of δ is deduced, in general, in a similar fashion to α' and γ , by deducting the absorption caused by the aliphatic groups ($-\text{CH}_3$, $>\text{CH}_2$) in the molecule and dividing the sum of the remainder by the sum of the $>\text{CH}$ (aromatic) bonds.

At all wave numbers, values of δ and δ' using α and α' were computed. The selection of the value for the unit coefficient of the $>\text{CH}$ (aromatic) bond was based on the randomness of the deviations, the value having the least trend in the residuals being considered superior. However, α' generally gave the best fit. The points at which α was used to develop δ are indicated in table 5. The effect of neglecting the unit coefficient δ (i. e., assuming that the aromatic nucleus did not contribute to the absorption) was tried at several wave numbers far removed from the maximum for aromatic absorption. Finally, the value of δ selected was that one conforming most closely with the observed data. In table 3 are given some results of using both α and α' on the series of aromatic hydrocarbons. Incidentally, the naphthenes shown were not included in the derivation of the coefficient, β .

TABLE 3.—Results of representing, for the cycloparaffins and aromatic hydrocarbons, the total molal absorption coefficient, as a function of the number n_a , of primary ($-\text{CH}_3$), n_b , of secondary ($>\text{CH}_2$) and, n_d , of $>\text{CH}$ (aromatic) groups (unrevised coefficients also given)

$$K = n_a \alpha' + n_b \beta + n_d \delta$$

Compound	Number of groups				8750 cm^{-1}					5900 cm^{-1} (unrevised)					5900 cm^{-1}				
	$-\text{CH}_3$	$>\text{CH}_2$	$>\text{CH}$	Arom.	$\alpha = 0.00327^a \beta = 0.00126 \delta = 0.01121$					$\alpha = 0.1574 \beta = 0.0057 \delta = 0.0651$					$\alpha' = 0.1256 \beta = 0.0057 \delta' = 0.0764$				
	n_a	n_b	n_c	n_d	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$	$\frac{\Delta}{100}$	$\frac{\Delta}{K_{\text{obs.}}} = \text{Number of } \frac{\delta}{\delta \text{CH (arom.)}} \text{ gr.}$	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$	$\frac{\Delta}{100}$	$\frac{\Delta}{K_{\text{obs.}}} = \text{Number of } \frac{\delta}{\delta \text{CH (arom.)}} \text{ gr.}$	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$	$\frac{\Delta}{100}$	$\frac{\Delta}{K_{\text{obs.}}} = \text{Number of } \frac{\delta}{\delta \text{CH (arom.)}} \text{ gr.}$
Benzene.....	0	0	0	6	0.0692	0.0672	+0.0020	+2.9	+0.2	0.516	0.391	+0.125	+23.2	+1.9	0.516	0.458	+0.058	+11.2	+0.8
Toluene.....	1	0	0	5	.0670	.0593	+.0077	+10.0	+.3	.541	.483	+.058	+10.7	+.9	.541	.508	+.033	+5.1	+.1
Ethylbenzene.....	1	1	0	0	.0634	.0606	+.0028	+4.6	+.2	.524	.489	+.035	+8.7	+.5	.524	.514	+.010	+1.9	+.1
n-Propylbenzene.....	1	2	0	0	.0702	.0618	+.0084	+11.9	+.4	.507	.495	+.012	+23.7	+.2	.507	.519	-.012	-2.4	-.2
Isopropylbenzene.....	2	0	1	5	.0675	.0625	+.0050	+7.4	+.4	.569	.641	-.072	-13.7	-1.1	.569	.633	-.064	-11.2	-.3
m-Xylene.....	2	0	0	4	.0390	.0513	-.0123	-32.4	-1.1	.443	.576	-.133	-28.6	-2.0	.443	.557	-.109	-24.3	-1.4
o-Xylene.....	2	0	0	4	.0551	.0513	+.0038	+5.9	+.3	.570	.576	-.006	-1.1	-0.1	.570	.557	-.013	-2.3	-0.1
p-Xylene.....	2	0	0	4	.0482	.0513	-.0031	-9.4	-0.3	.673	.576	+.102	+15.1	+1.6	.673	.557	+.121	+17.9	+1.6
Mesitylene.....	3	0	0	3	.0297	.0434	-.0137	-46.2	-1.3	.516	.667	-.151	-29.2	-2.3	.516	.606	-.090	-17.5	-1.2
Hemimellitene.....	3	0	0	3	.0494	.0434	+.0060	+12.2	+0.5	.601	.667	-.066	-11.0	-1.0	.601	.606	-.005	-0.8	-0.1
Pseudocumene.....	3	0	0	3	.0372	.0434	-.0062	-16.7	-0.6	.605	.667	-.062	-10.2	-1.0	.605	.606	-.001	-0.2	-0.0
Average deviation.....	-----	-----	-----	-----	-----	-----	-----	±13.5	±0.6	-----	-----	-----	±15.7	±1.1	-----	-----	-----	±8.7	±0.6

^a $\alpha = \alpha'$ at this point.

Compound	Number of groups			5781 cm ⁻¹ (unrevised)					5781 cm ⁻¹				
	-CH ₃	>CH ₂	≧CH	$\alpha=0.1008 \beta=0.0908$					$\alpha'=0.0708 \beta=0.0908$				
	n_a	n_b	n_c	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \times 100$	$\frac{\Delta}{\beta} = \text{Number of } >\text{CH}_2 \text{ gr.}$	$K_{\text{obs.}}$	$K_{\text{calc.}}$	$K_{\text{obs.}} - K_{\text{calc.}}$ Δ	$\frac{\Delta}{K_{\text{obs.}}} \times 100$	$\frac{\Delta}{\beta} = \text{Number of } >\text{CH}_2 \text{ gr.}$
Cyclohexane.....	0	6	0	0.736	0.545	+0.191	+25.9	+2.1	0.736	0.545	+0.191	+25.9	+2.1
Methylcyclohexane.....	1	5	1	.662	.555	+0.107	+16.2	+1.2	.662	.525	+0.137	+20.7	+1.5
Ethylcyclohexane.....	1	6	1	.623	.645	-0.023	-3.7	-0.3	.623	.616	+0.007	+1.1	+0.1
Average deviation.....							±15.3	±1.2				+15.9	+1.2
1, 2-Dimethylcyclopentane (bp 99.3°).....	2	3	2	.363	.474	-.111	-30.6	-1.2	.363	.414	-.051	-14.2	-0.6
1, 2-Dimethylcyclopentane (bp 91.8°).....	2	3	2	.358	.474	-.116	-32.4	-1.3	.358	.414	-.056	-15.7	-.6
Methylcyclopentane.....	1	4	1	.274	.464	-.190	-69.4	-2.1	.274	.434	-.160	-58.4	-1.6
Average deviation.....							-44.1	-1.5				-29.4	-1.0

IV. RESULTS

The results, listed under the names of the compounds, are summarized in table 4. The formulas and numbers of structural units comprising the molecule are listed in appropriate columns. The values given in this table in the columns headed primary, secondary, etc., indicate the deviations between calculated and observed ordinates of the molal absorption curves in terms of the structural units of particular interest at the wave number under which these values appear. The average deviation is shown for each series of compounds (such as isoparaffins, cycloparaffins, etc.) and is calculated as the arithmetical mean of the individual deviations. Values in parentheses have been omitted from the averages by assuming that they are mistakes.

The unit coefficients of the several groups (α' , β , γ , and δ) are found in table 5, where they are arranged in order of increasing wave number.

TABLE 4.—Summary of deviations of calculated values of the molar absorption coefficient from the observed values; these residuals being expressed in terms of group sought

Compound	Formula	Number of groups				Primary $-\text{CH}_3$							Secondary $>\text{CH}_2$						Tertiary $\geq\text{CH}$		Aromatic ($>\text{CH}$)		
		$-\text{CH}_3$	$>\text{CH}_2$	$\geq\text{CH}$	Arom.	45800	5860	5900	6975	7350	8400	8680	5668	5781	7064	7185	8254	8300	6950	8160	5900	8680	8750
2-Methylbutane.....	C_5H_{12}	n_a	n_b	n_c	n_d	-0.2	-0.2	-0.3	-0.8	-0.3	0.0	-0.4	0.0	+0.3	-0.7	+0.3	+0.1	+0.2	-1.5	+0.1	-0.5	-0.5	-0.2
2-Methylpentane.....	C_6H_{14}	1	1	1	1	+4	+6	+3	+2	-3	+5	+2	+1	+2	-1	-4	+2	+3	+8	+3	+3	+2	-2
2-Methylhexane.....	C_7H_{16}	3	3	3	3	+2	+2	+2	+4	+1	+4	+5	0	+2	+1.2	+9	+3	+7	+2	+2	+1.2	+5	+2
2-Methylheptane.....	C_8H_{18}	3	3	3	3	+3	+2	+1	0	-3	+3	0	+3	-6	+2	-3	+3	+7	0	+1.1	+1	0	-3
3-Methylpentane.....	C_6H_{14}	3	3	3	3	+8	+3	-1	9	-1	+4	+4	+1.4	+9	+1.0	+7	+8	+1.2	-2.1	+7	-2	+5	+1
3-Methylhexane.....	C_7H_{16}	3	3	3	3	+4	+3	+1	+1	+1	+4	-1	0	+1	-7	+5	-1	-1.5	-2	-2	+2	-1	+1
3-Methylheptane.....	C_8H_{18}	3	3	3	3	+1.0	+4	+6	-2	-2	+3	+1	+2	0	-3	-1	-2	+4	-9	+1	+1.0	+1	0
4-Methylpentane.....	C_6H_{14}	3	3	3	3	-1	+2	+4	-1	0	+4	+8	-6	-1.1	-5	-1	-2	+5	-1.0	+4	+6	+8	-1
4-Methylhexane.....	C_7H_{16}	3	3	3	3	-5	+4	+3	-2	-5	-3	-1.4	-2	-9	0	-2	-6	-5	-4	+4	+5	-1.4	-1.1
5-Methylpentane.....	C_6H_{14}	3	3	3	3	+3	+3	0	-7	-1	+7	+4	-2	-4	-1.3	+3	-5	+2	-2.0	-1.2	0	+4	+3
3,3-Dimethylpentane.....	C_7H_{16}	4	2	0	0	-9	-3	0	-4	-2	0	-3	-3	-4	-8	0	-7	-1.1	+1.0	-7	0	-4	0
2,3-Dimethylbutane.....	C_6H_{14}	4	0	2	2	+4	-2	-7	(+2.2)	-7	+1	-2	(+2.8)	+1.1	+1.1	+1	+1.7	+1.8	+2.2	+8	-1.2	-2	-1
2,3-Dimethylpentane.....	C_7H_{16}	4	1	2	2	0	-1	+3	+4	+4	0	-1	+1	+3	+3	+1	+6	-1	-3	+6	+5	-1	+2
2,4-Dimethylpentane.....	C_7H_{16}	4	1	2	2	-2	+3	+1	+5	-2	-2	-3	+1	+4	+4	-1	-2	-1.0	-2	-7	+2	-3	+2
2,2,3-Trimethylbutane.....	C_7H_{16}	5	0	1	1	-1.3	(-2.2)	-6	-4	-1	-6	0	-1.4	-1	-1.7	-1.4	-6	-1.0	+1.1	-1	-1.0	0	+4
2,2,4-Trimethylpentane.....	C_8H_{18}	5	1	1	1	-1.7	-7	-1.1	0	+6	-3	+4	-1.6	-7	-1	-2	+1	-3	+1.5	-2	-1.8	+4	+2
3,5,5-Trimethylheptane.....	$\text{C}_{10}\text{H}_{22}$	5	3	1	1	-9	+1	+9	-1.2	+7	-2	+3	-7	-4	-1.7	-2	-1.3	-1.4	-1.1	-1.0	-1.5	+3	+3
3,4,5,5-Tetramethylhexane.....	$\text{C}_{10}\text{H}_{22}$	6	1	2	2	-1.1	-5	-2	-5	+5	-7	0	-7	-1	-1.3	-2	-2	-5	-4	-1.1	-4	0	+3
Average deviation.....						± 6	± 3	± 4	± 4	± 3	± 3	± 3	± 5	± 5	± 7	± 4	± 5	± 8	± 1.0	± 5	± 6	± 3	± 2
Cyclohexane.....	C_6H_{12}	0	6	0	0	+1.6	+2	+2	-6	-4	+1.0	-1	-2	+2.1	-8	+1	-2	+2.0	-1.0	-2	+4	-1	-3
Methylcyclohexane.....	C_7H_{14}	1	5	1	1	+5	-3	-1	+1	-2	+3	-3	+1.2	+1.5	-1.5	+2	-2	+1.2	-8	+3	-1	-3	-4
Ethylcyclohexane.....	C_8H_{16}	1	6	1	1	+1.2	+4	+7	-6	-4	+5	-4	-9	+1	-4	+2	-5	+9	-9	+2	+1.2	-4	-3
Average deviation.....						+1.1	± 3	± 3	± 4	-3	+6	-3	± 8	+1.2	-9	+2	-3	+1.4	-1.2	± 2	± 6	-3	-3
1,2-Dimethylcyclopentane (bp. 99.3°).....	C_7H_{14}	2	3	2	2	-2	+4	+1.0	-7	0	+8	-4	+8	-6	-1.4	-1.6	-8	0	-2.1	-1.4	+1.6	-5	-1
1,2-Dimethylcyclopentane (bp. 91.8°).....	C_7H_{14}	2	3	2	2	0	+1	+7	-2	-1	+4	-2	-5	-6	-2	-1.2	-1.2	-9	-1.7	-2.0	+1.1	-2	-4
Methylcyclopentane.....	C_6H_{12}	1	4	1	1	-3	+9	+1.4	-5	+7	+1.2	-1.6	-1	-1.8	-7	-1.4	-1.4	-5	-1.5	-1.5	+2.3	-1.5	+2
Average deviation.....						-2	+5	+1.0	-5	± 3	+8	-3	± 5	-1.0	-8	-1.4	-1.1	-5	-1.8	-1.6	+1.7	-7	± 2

* Wave number values in cm^{-1}

TABLE 4.—Summary of deviations of calculated values of the molar absorption coefficient from the observed values; these residuals being expressed in terms of group sought—Continued

Compound	Formula	Number of groups				Primary —CH ₃								Secondary >CH ₂						Tertiary >CH		Aromatic (>CH)		
		—CH ₃	>CH ₂	>CH	Arom.	5800	5860	5900	6975	7350	8400	8680	5668	5781	7064	7185	8254	8300	6950	8160	5900	8680	8750	
Benzene.....	C ₆ H ₆	0	0	0	6	—3	+2	+5	+8	—3	+1	—3.2	—9	—2	.0	+7	+2	+1	.0	+2	+8	—3.4	+2	
Toluene.....	C ₇ H ₈	1	0	0	5	+2.3	+4	+3	+5	.0	+1	—8	—1	+5	+5	+4	.0	+1	+1.1	+1	+4	—8	+6	
Ethylbenzene.....	C ₈ H ₁₀	1	1	0	5	—3	—2	+1	—4	—6	+7	—6	+2	—8	—5	—5	—3	+5	.0	+1	—2.1	+3		
n-Propylbenzene.....	C ₉ H ₁₂	1	2	0	5	—3	—2	—1	—4	—6	+7	—6	—7	—9	—5	—5	—1	+7	—1.0	—2	—2	+6	+8	
Isopropylbenzene.....	C ₉ H ₁₂	2	0	0	5	—6	—9	—5	—3	—3	+4	—3	+1	—8	—2	+1	—4	+4	—1.1	+1	—8	—3	+4	
m-Xylene.....	C ₈ H ₁₀	2	0	0	4	—6	—9	—9	.0	.0	.0	—1.0	—7	.0	+4	—3	.0	.0	+3	+2	—1.4	—1.1	—1.1	
o-Xylene.....	C ₈ H ₁₀	2	0	0	4	+3	—1	—1	+1	—4	—1	—4	+2	+8	+1	—3	.0	—1	+1	+1	—1	—4	+3	
p-Xylene.....	C ₈ H ₁₀	2	0	0	4	+1	+3	+1.0	—1	.0	+3	+2.7	+1	+5	+1	+4	—1	+3	+2	+3	+1.6	+2.9	—3	
Mesitylene.....	C ₉ H ₁₂	3	0	0	3	—2	+6	—7	—7	+4	+4	+2.2	+9	—1.2	.0	.0	—1	+4	—2	+3	—1.2	+2.4	—1.2	
Hemimellitene.....	C ₉ H ₁₂	3	0	0	3	—5	—6	.0	—3	+3	—6	+3	+5	+3	+4	—7	+2	—6	—5	+3	—1	+4	+5	
Pseudocumene.....	C ₉ H ₁₂	3	0	0	3	—1	+1	.0	—6	—5	—1	+1.3	+9	+9	—1	—1	+3	—1	+3	+4	.0	+1.4	—6	
Average deviation.....						±.5	±.4	±.4	±.4	±.3	±.3	±1.3	±.5	±.6	±.2	±.4	±.2	±.3	±.5	±.2	±.6	±1.4	±.6	
1-Phenyl-n-octadecane.....	C ₂₄ H ₄₂	1	17	0	5	+1.3	+4	+4	—6	—1.5	+1	—2.0	+6	+3	—3	—1.1	—1.3	—3	—1.6	—9	+1.0	—2.6	—1.8	
1-Cyclohexyl-n-octadecane.....	C ₂₄ H ₄₈	1	22	1	0	+1.6	+6	+6	—9	—1.5	+5	—1.5	—3	+1.3	—4	—2.1	—1.8	+4	—2.8	+8	+1.4	—2.0	—1.3	
5-Cyclohexyl-n-docosane.....	C ₂₈ H ₅₆	2	24	2	0	+2.5	+6	+6	—1	—1.4	+5	—1.3	—2.3	—3	+3.7	—1.6	—1.9	+3	—3.9	—1.0	+1.5	—1.7	—1.0	
1-(7-Tetrahydronaphthyl)-n-octadecane.....	C ₂₈ H ₄₈	1	21	0	3	+1.6	+7	+7	—8	—1.4	+8	+7	+3.7	—2	+1.1	—1.9	—1.4	+6	—1.8	+2.3	+1.8	+9	—1.9	
5-(7-Tetrahydronaphthyl)-n-docosane.....	C ₃₂ H ₅₆	2	23	1	3	+1.5	+6	+6	—1.2	—1.5	+1.2	+8	+1.3	—1	—1.1	—1.1	—1.0	+1.3	—3.8	—2	+1.6	+1.1	—2.4	
5-(7-Tetrahydronaphthyl)-n-docosene-5.....	C ₃₂ H ₅₄	2	22	(1)	3	+1.8	+9	+9	.0	—1.0	+1.0	+2	+2.0	—1.2	+5	—1.4	—1.7	+7	—1.1	+6	+2.3	+2	—2.1	
5-(2-Decahydronaphthyl)-n-docosane.....	C ₃₂ H ₅₆	2	26	4	0	+3.8	+5	+5	—6	—3.0	+8	—1.4	+4.1	(+5.8)	+7	—9	+1	+2.4	—5.3	+4	+1.1	—1.9	—1.4	
1-(p-Diphenyl)-n-octadecane.....	C ₃₀ H ₄₆	1	17	0	9	+1.6	—1.0	—1.0	—7	—1.6	+3	—2.0	—4	+2	+7	—1.0	—9	—2	—2.3	+2	—2.4	—2.6	—3	
5-(p-Diphenyl)-n-docosene-5.....	C ₃₄ H ₅₂	2	18	(1)	9	+7	—9	—9	(+4.2)	—9	+2	—2.4	—1.3	—9	(+11.1)	+7	—1.7	—6	(+6.1)	+4	—2.1	—3.1	+4	
Average deviation.....						+1.8	±.7	±.7	—6	—1.5	+5	±1.4	±1.8	±.6	±1.1	1.3	1.3	±.8	—2.8	±.8	±1.7	±1.8	±1.4	

• Wave number values in cm⁻¹

TABLE 5.—Values of unit coefficients of the various groups at the indicated wave numbers

Wave number	Group sought	$-\text{CH}_3$	$>\text{CH}_2$	$\geq\text{CH}$	$>\text{CH}$ (aromatic)
cm^{-1}		α'	β	γ	δ
5668	$>\text{CH}_2$ -----	0.0529	0.0638	-----	0.0216
5781	$>\text{CH}_2$ -----	.0708	.0908	-----	.0137
5800	$-\text{CH}_3$ -----	.1052	.0674	-----	.0138
5860	$-\text{CH}_3$ -----	.1363	.0194	-----	^b .0314
5870	$-\text{CH}_3$ -----	.1426	.0130	-----	.0376
5900	$-\text{CH}_3, >\text{CH}$ (aromatic)-----	.1256	.0057	-----	.0764
6950	$>\text{CH}$ -----	.00718	.00748	0.00532	.00175
6975	$-\text{CH}_3$ -----	.00948	.00776	-----	^b .00136
7064	$>\text{CH}_2$ -----	.01253	.00956	-----	.00134
7185	$>\text{CH}_2$ -----	.01791	.00978	-----	.00118
7350	$-\text{CH}_3$ -----	.01025	.00164	-----	.00184
8160	$>\text{CH}$ -----	.00197	.00092	.01132	-----
8254	$>\text{CH}_2$ -----	.00629	^a .01989	-----	-----
8300	$>\text{CH}_2, -\text{CH}_3$ -----	.00948	.01585	-----	-----
8400	$-\text{CH}_3$ -----	.02705	.00735	-----	-----
8680	$-\text{CH}_3$ -----	.00663	^a .00176	-----	.00622
8750	$>\text{CH}$ (aromatic)-----	.00327	.00126	-----	.01121

^a Developed using C_8 to C_{20} (omitting C_{10}).^b Developed by using α .

The method used in table 4 for expressing the correspondence between the observed and calculated data is best suited to the present purpose. The principal factor that fixes the accuracy is the percentage transmission. Since there are optimum values for the ratio I/I_0 , the concentration of a compound under investigation is adjusted so that its maximum of absorption lies within certain limits. Consequently, a homologous series such as the normal paraffins, for which the maximum does not shift appreciably, has the same percentage error for the various members at any specific wave number. However, the error for one compound at different wave numbers within a single region is more nearly absolute than relative, and comparisons at the same wave number between compounds having maxima at different wave numbers can not be made on a basis of percentage error.

It should be noted that several effects combine to give less accuracy in the regions 5400 to 6400 cm^{-1} and 6400 to 7400 cm^{-1} than in the region 7900 to 8900 cm^{-1} . These are as follows: (1) The steepness of the curves that causes large errors if, when reading the photographic plate of the automatically recorded transmission curve, the proper position with respect to the fiducial point is not maintained; (2) the greater complexity of the regions 5400 to 6400 cm^{-1} and 6400 to 7400 cm^{-1} ; (3) the proximity of the points at which maxima of absorption for the various groups appear; and (4) the absorption of atmospheric water in the region 6400 to 7400 cm^{-1} .

TABLE 6.—*List of wave numbers*

[Arranged according to groups for which used; with comments.]

Group sought	Wave number	Remarks
—CH ₃	8400	Maximum for —CH ₃ ; best point.
—CH ₃ , >CH ₂	8300	Excellent check point.
—CH ₃	5900	Maximum for —CH ₃ ; strong absorption by >CH (aromatic).
—CH ₃	5880	Maximum for —CH ₃ .
—CH ₃	5870	Check point.
—CH ₃	6975	Do.
—CH ₃	7350	Do.
>CH ₂	8254	Maximum for >CH ₂ ; best point.
>CH ₂	5781	Maximum for >CH ₂ .
>CH ₂	5668	Naphthenes different at this point.
>CH ₂	5800	Do.
>CH ₂	7185	Check point.
>CH ₂	7075	Affected by substitution; check point.
>CH.....	8160	Maximum for >CH; best point.
>CH.....	6950	Check point.
>CH (aromatic).....	8750	Maximum for >CH (aromatic); best point.
>CH (aromatic).....	8880	Check point. Second maximum of —CH ₃ .
>CH (aromatic).....	5900	Check point; strong absorption by —CH ₃ .

The region 5,400 to 6,400 cm^{-1} does have the advantage that the proper concentration is one-fifth that of the other regions. This reduces the opportunity for departures from Beer's law [31, 34].

In table 6 the wave numbers are listed in order of importance with respect to the various groups, together with appropriate remarks concerning conditions at that particular wave number.

V. DISCUSSION

1. NORMAL PARAFFINS

The normal paraffins show, in general, excellent agreement with the calculated results (see table 1). The average deviations are less than one structural group. However, the lower members (*n*-pentane to *n*-heptane) exhibit, in most cases, an anomalous behavior. The occurrence of larger deviations in the case of the smaller molecules rather than of the larger ones is to be expected because the latter are composed mainly of "average" $>\text{CH}_2$ groups, in the sense that these units may be considered equivalent to each other and unaffected by lengthening of the chain (i. e., the intramolecular environment is more constant on the average). The remainder of the variations, excepting those for *n*-decane, are well within the experimental precision. No alternating effects (caused by an odd or even number of carbon atoms in the chain) are noticeable.

2. ISOPARAFFINS

The data for the isoparaffins do not agree, in general, as well as those for the normal paraffins. The values appearing in table 5 as α' were used in all cases tabulated in table 4. However, comparison of these values with the values of α listed in section III reveals that at 7064, 7350, 8254, and 8750 cm^{-1} , α' is identical with α . In these instances, the algebraic sums of the residuals resulting from the application to the isoparaffins of the coefficients, α and β , developed for the normal paraffins were almost zero. The equality of α and α'

at these points may be explained either by saying that the decrease usually encountered in α by passing from a normal paraffin to an isoparaffin is offset by an increase in the coefficient for some other group (see next paragraph), or that the mode of vibration associated with absorption at these wave lengths is unchanged when the methyl ($-\text{CH}_3$) group is shifted from the end of the chain to some intermediate position.^{16a}

The true hybrid nature of α' should be emphasized. Some of the methyl groups used in its derivation probably have the same value as those in the normal paraffins. It is reasonable to suppose that the single methyl group at the extremity of 2-methylnonane is the same as a methyl group in *n*-nonane. Actually, a differentiation can be made between "normal" methyl groups ($-\text{CH}_2-\text{CH}_3$) and "iso" methyl groups ($-\text{CH} \begin{smallmatrix} \text{CH}_3 \\ \text{CH}_3 \end{smallmatrix}$). This effect may explain the fact that there is a negative trend of the residuals in the revised data on the isoparaffins as the proportion of methyl groups increases, showing that the average coefficient is too great when the ratio of "iso" to "normal" methyl groups becomes large. Thus, strictly, there can be at least two types of methyl groups with their respective coefficients.¹⁷

Practically, the simplification attained by the retention of a single average coefficient for any group is necessary where unknowns are to be analyzed. Due regard for the deviations in the data on similar structures previously studied permits a comparison to be made between an unknown compound of low molecular weight and an assumed structure. The author believes that as additional data are accumulated distinctions can logically be made between the types of methyl groups. However, the present data do not furnish justification for these refinements. The series of methylnonanes should furnish opportunity for the discrimination between the two types of methyl groups. For example, 2-methylnonane contains two "iso" and one "normal" methyl groups, while 5-methylnonane has one "iso" and two "normal" methyl groups. The other isomers are intermediate and conceivably might be a means of evaluating the distance over which interaction between two methyl groups can be exerted. Inspection of table 4 conclusively shows that the variations between the methylnonanes are of the order of the experimental error.

Theoretical considerations would also demand an investigation concerning the change of value for $>\text{CH}_2$ groups with respect to the relative position of the nearest $-\text{CH}_3$ group. The method employed for the derivation of α' takes no cognizance of such corresponding possible variations in β but imposes the effects on α' . Moreover, included in α' is the absorption, if any, due to $>\text{CH}$ groups (except at 6950 and 8160 cm^{-1}). Thus, owing to the structure of the majority of the compounds studied, α' may represent more nearly the value

for the group $\text{HC}-\text{CH}_3$. The presence of the various types of methyl groups will modify the composite value of α' .

The analysis for tertiary ($>\text{CH}$) groups is poor in contrast with the other groups. This is to be expected since the group contains only one $\text{C}-\text{H}$ bond, and its absorption is therefore relatively weak.

^{16a} See Wulf [37] for a discussion of this point.

¹⁷ A third type would be that in which two methyl groups are joined to the same carbon atom of a chain. Minor effects might also be noticed due to the proximity of other methyl groups (i. e., a 2,3-dimethyl compound).

Even at its maximum, the majority of the absorption is due to other groups and inaccuracies in the total molal absorption coefficient cause considerable errors in the analysis for tertiary groups. Fair results are obtained at the wave number 8160 cm^{-1} .

3. NAPHTHENE AND AROMATIC HYDROCARBONS

The results for the naphthenes indicate a slight departure from the values of the unit coefficients developed for the paraffin hydrocarbons. It has been previously noted by Liddel and Kasper [2] that naphthenic secondary ($>\text{CH}_2$) groups are not exactly equivalent to similar paraffinic groups.

A critical source of error in the results for the naphthenic as well as the aromatic hydrocarbons lies in the inadequate number of compounds available for study. Unsubstituted rings, such as benzene and cyclohexane, do not necessarily furnish a basis for unit coefficients, since substitution, by destroying their symmetry, may lead to different behavior. Furthermore, the unit coefficient for a $-\text{CH}_3$ group joined directly to the ring is not necessarily the same as that found in aliphatic compounds. This obviates the possibility of any study of the ring if the absorption of the $-\text{CH}_3$ groups is proportionately large. It was observed in a previous paper [18] and substantiated in this paper (see table 3) that the mono-substituted alkyl benzene derivatives have absorption spectra that are more constant as to position and magnitude of the maxima arising from the aromatic bonds than do the polysubstituted derivatives.¹⁸ Consequently, data are needed on a greater number of mono-substituted derivatives of both the naphthenes and aromatics before the analysis of such structures can be as satisfactory as that for the paraffin hydrocarbons.

An additional factor complicating the investigation of the polysubstituted derivatives of the cyclic hydrocarbons is the dissimilarity between the spectra of isomers obtained by varying the positions of the substituents on the nucleus. The existence of interaction with mutual alterations may be demonstrated by studying the absorption spectra of compounds in regions where either the nucleus or the substituent groups do not absorb. Thus Barnes [40] found that the absorption of benzene is modified not only by the replacement of a hydrogen atom by halogen but also by a variation in halogen (chlorine to bromine). Wulf and coworkers [36, 41], in a study of OH and NH bonds, have recorded absorption spectra which vary with the spatial distribution of the absorbing groups on the inactive nucleus (i. e., catechol, resorcinol, and hydroquinone; the *o*-, *m*-, and *p*-phenylenediamines; and the benzidines). A theoretical explanation for such effects has been furnished by Pauling [43]. No doubt similar effects of lesser magnitude are manifested in the cyclic hydrocarbons and influence the absorption of both the nucleus and the substituent groups.

Notwithstanding such theoretical possibilities, the criterion for the validity of the method given in this paper lies in the degree of concordance obtained by the practical application. Inspection of table 4

¹⁸ This fact has also been observed by Barnes [40] for the region of the fundamental. Using the very high resolving power of an echelette grating, he recorded the percentage transmission of benzene and some of its derivatives for the region 3.0 to $3.6\text{ }\mu$. Only a small shift ($0.007\text{ }\mu$) was noted in the benzene bands of mono-substituted alkyl derivatives while the relative intensities of these bands were not greatly changed. The xylenes, however, showed dissimilarities not only from benzene but among themselves.

reveals that a correct analysis for most of the aromatic hydrocarbons and a nearly correct analysis for the naphthenes can be expected after careful selection of the proper wave lengths. Some uncertainty is, however, attached to the method in the case of naphthenes of low molecular weight. Further data on longer chain derivatives would, the author believes, permit removal of the present discrepancies.

4. HYDROCARBONS OF HIGH MOLECULAR WEIGHT

The hydrocarbons of high molecular weight are an excellent test of the agreement of experimental with calculated molal absorption coefficients. Since these compounds were not used in the development of the unit coefficients, the agreement of the values cannot be ascribed to the manner in which the constants were evaluated. From table 4 it is apparent that the number of groups is accurate in the more favorable instances to ± 1 group. The average percentage deviations (not shown) are generally ± 8 percent, which is roughly the experimental error. A more exact analysis of such large molecules can be expected with improved experimental technique.

The hydrocarbons studied include several structural groups concerning which there is no available information from the present data. The olefin bonds in two of the compounds are of minor importance as only one C—H is involved, unless the unsaturated bond affects adjacent groups. These bonds are indicated in the column for $>\text{CH}$ groups and are italicized since they were neglected in the computations. As observed in the previous paper [18], the C—H bonds in the benzenoid ring of tetrahydronaphthalene are not necessarily equivalent to similar bonds in benzene. The fact that the tetrahydronaphthyl derivatives alone show positive deviations at 8680 cm^{-1} , and, in general have higher negative deviations at 8750 cm^{-1} , substantiates the hypothesis [18] that the maximum of absorption for the "aromatic" portion of the tetrahydronaphthyl nucleus has shifted to lower wave numbers.

The consistently negative deviations for almost all the compounds at 7185 and 8254 cm^{-1} can be explained as a modification of the absorption coefficients for the $>\text{CH}_2$ groups,¹⁹ owing to the proximity of the cyclic nuclei. Although these deviations are still within the experimental error, the fact that they are not random is significant. Their virtually constant value precludes the possibility of their dependence upon structural variables other than one common to all.

At the wave numbers 5668 and 5800 cm^{-1} , the deviations of the partially and totally hydrogenated ring structures can be correlated with the number of secondary ($>\text{CH}_2$) groups included in the ring. Possibly in the other regions there are similar wave numbers at which data were not calculated. This behavior might furnish a method for distinguishing between paraffinic and naphthenic secondary groups. It would, however, first be necessary to establish that the two types of bonding are equivalent at one wave length and different at another.

¹⁹ See footnote 8, p. 132.

VI. DETERMINATION OF THE NUMBER AND KIND OF STRUCTURAL GROUPS IN UNKNOWN HYDROCARBONS

1. GENERAL PROCEDURE

In utilizing the present analytical method to determine the number and kind of structural groups in unknown hydrocarbons, the following procedure is used:

1. Measurement is made of the transmission of the given hydrocarbon in solution in carbon tetrachloride in the regions 6400 to 7400 cm^{-1} , 7900 to 8900 cm^{-1} , and 5400 to 6400 cm^{-1} , with the concentration so adjusted that the transmission lies in the limits 40 to 60 percent [2, 34].

2. The transmission curves are reduced to curves of total molal absorptive index, K , [2, 34].

3. For the given wave numbers selected for determining the various structural groups, values of the total molal absorptive index, K , are recorded.

4. Combination of these experimental values of K with the values of the unit coefficients, α' , β , γ , δ at the appropriate wave numbers, yields a series of simultaneous equations, of the general form,

$$K = n_a \alpha' + n_b \beta + n_c \gamma + n_d \delta \quad (8)$$

5. These simultaneous equations are solved to obtain values for n_a , n_b , n_c , and n_d , the number of $-\text{CH}_3$, $>\text{CH}_2$, $\geq\text{CH}$, and aromatic groups respectively. To each value so obtained an estimated uncertainty is assigned.

6. These values of the various structural groups are checked by calculating the total molal absorption index at wave numbers where the use of simultaneous equations is not feasible because of physical indeterminacy.

In practice, a simplification of rule 5 may be effected. A visual inspection can be made of the molal absorption curve to determine if absorption due to aromatic bonds in the molecule is evident. In the absence of significant absorption in the region near 8750 cm^{-1} , the series of simultaneous equations except those for 8160 and 6950 cm^{-1} may be reduced to two unknowns, the number of methyl ($-\text{CH}_3$) groups, n_a , and of methylene ($>\text{CH}_2$) groups, n_b . After properly evaluating n_a and n_b , the number of tertiary ($\geq\text{CH}$) groups may be calculated from data on K at 8160 and 6950 cm^{-1} . The assumed absence of aromatic bonds in the molecule should be verified by a comparison at 8750 cm^{-1} of the observed total molal absorption coefficient with the value calculated from the various aliphatic bonds.

When, however, "aromatic" absorption appears to be present, a series of simultaneous equations involving α' , β , γ , and δ should be set up and solved.

The accuracy of the final result is determined by assigning to the total molal absorption index an error of ± 8 percent. This error is carried through the equations in absolute value and finally converted into terms of the group sought.

When all the desired values for n_a , n_b , n_c , and n_d have been calculated, they are carefully considered in connection with the following

expressions, which can be written from a knowledge of the molecular weight and the carbon-hydrogen content of the unknown hydrocarbon:

$$\text{C}_n\text{H}_{(2n+x)} \quad (9)$$

$$n_a + n_b + n_c + n_d \leq n \quad (10)$$

$$3n_a + 2n_b + n_c + n_d = 2n + x \quad (11)$$

If there are present carbon atoms to which no hydrogen atoms are joined, the total number of carbon atoms calculated from the infrared spectra will be less than the number, n , from the molecular weight. Obviously, the total number of carbon atoms can never be greater than the number fixed by the molecular weight. Under any conditions, eq 11 must hold. These equations are essentially "material" balances, eq 10 relating to carbon and eq 11 to hydrogen.

2. EXAMPLE: DETERMINATION OF THE STRUCTURE OF AN UNKNOWN ISOMERIC NONANE

The foregoing procedure is illustrated by its application to the determination of the number and kind of structural groups in an unknown "isononane" which was recently isolated from a midcontinent petroleum by White and Rose [42].

Figure 2 shows the curves of the molal absorptive index for the given unknown "isononane" for the regions

5400 to 6400 cm^{-1} , 6400 to 7400 cm^{-1} , and 7900 to 8900 cm^{-1} .

From the values of the molal absorptive index, K , obtained from these curves and values of the unit coefficients, α' , β , γ , given in table 5, the following system of equations may be written

$$\text{At } 5668 \text{ cm}^{-1} \quad 0.465 = 0.0529 n_a + 0.0638 n_b \quad (11a)$$

$$5781 \quad .608 = .0708 n_a + .0908 n_b \quad (11b)$$

$$5800 \quad .632 = .1052 n_a + .0674 n_b \quad (11c)$$

$$5860 \quad .599 = .1363 n_a + .0194 n_b \quad (11d)$$

$$5900 \quad .535 = .1256 n_a + .0057 n_b \quad (11e)$$

$$6975 \quad .0595 = .00948 n_a + .00776 n_b \quad (11f)$$

$$7064 \quad .0793 = .01253 n_a + .00956 n_b \quad (11g)$$

$$7185 \quad .0960 = .01791 n_a + .00978 n_b \quad (11h)$$

$$7350 \quad .0411 = .01025 n_a + .00164 n_b \quad (11i)$$

$$8254 \quad .0808 = .00629 n_a + .01989 n_b \quad (11j)$$

$$8400 \quad .1230 = .02705 n_a + .00735 n_b \quad (11k)$$

In solving, the equations have been paired as follows for methyl ($-\text{CH}_3$): 11d, 11e; 11e, 11a; 11e, 11b; 11f, 11k; 11j, 11k; 11c, 11d; 11c, 11i; 11f, 11i; for methylene ($>\text{CH}_2$) 11e, 11b; 11e, 11a; 11h, 11j; 11e, 11j; 11j, 11k; 11c, 11d; 11c, 11j. The above pairings were found to be the most favorable combinations of the unit coefficients to yield accurate values of the group sought. Other pairings either gave indeterminate solutions or included wave numbers where the analysis was not accurate.

In table 7 are given the results of the solution of the above simultaneous equations. The values of the groups are given in the order of decreasing reliability as judged by the error. The error is expressed in terms of unit groups and is found as explained in the preceding section. Under "wave numbers used" are listed the points at which data were taken to furnish the two equations whose solution yielded

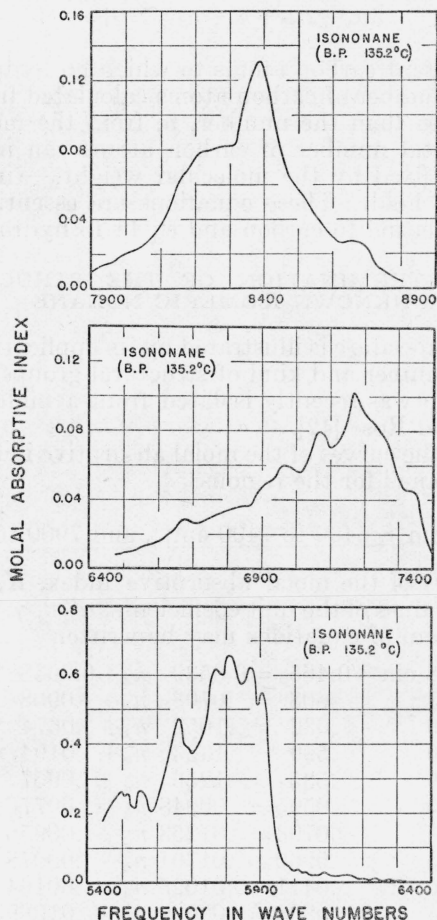


FIGURE 2.

the value shown. In only a few instances was a single pair of equations favorable to the determination of both methyl ($-\text{CH}_3$) and methylene ($>\text{CH}_2$) groups. Where indeterminate results were obtained, the values have been omitted from consideration. From table 7, the number of $-\text{CH}_3$ groups, n_a , is fixed as four and the number of $>\text{CH}_2$ groups as three.

In table 8, the same data are treated in a different manner. Where methyl ($-\text{CH}_3$) groups are to be determined, the integral value of methylene ($>\text{CH}_2$) groups is assumed (as determined from table 7) and vice versa. Both $-\text{CH}_3$ and $>\text{CH}_2$ groups are assumed to find the number of tertiary ($>\text{CH}$) groups.

TABLE 7.—Results of solution of the simultaneous equations 11a to 11k

Calculated number of groups	Wave numbers used
METHYL ($-\text{CH}_3$)	
4.10 \pm 0.35	5900, 5781
3.77 \pm .40	8254, 8400
4.20 \pm .50	5860, 5900
3.94 \pm .47	5800, 5860
3.69 \pm .55	6975, 8400
4.08 \pm .35	5900, 5668
3.35 \pm .51	5800, 7350
3.46 \pm .42	6975, 7350
METHYLENE ($>\text{CH}_2$)	
2.87 \pm 0.36	8254, 8400
2.76 \pm .33	8254, 5900
2.86 \pm .43	8254, 7185
2.71 \pm .45	8254, 5800
3.50 \pm .63	5900, 5781
3.91 \pm .67	5900, 5668
3.23 \pm 1.19	5800, 5860

TABLE 8.—Results of application of integral values of groups sought to data at all available wave numbers

Wave number	Molal absorptive index			Calculated number of groups
	$K_{\text{obs.}}$	$K_{\text{calc.}}$	% Error	
METHYL ($-\text{CH}_3$)				
5800	0.632	0.623	+1.4	4.1 ± 0.48
5860	.599	.603	-0.7	$4.0 \pm .35$
5900	.535	.520	+2.8	$4.1 \pm .34$
6975	.0595	.0612	-1.2	$3.9 \pm .50$
7350	.0411	.0459	-11.7	$3.5 \pm .32$
8400	.1230	.1302	-5.9	$3.8 \pm .36$
8680	.0263	.0318	-----	$(3.2) \pm (.6)$
METHYLENE ($>\text{CH}_2$)				
5668	0.465	0.403	+13.3	4.0 ± 0.58
5781	.608	.556	+8.6	$3.6 \pm .54$
7064	.0793	.0788	+0.6	$3.1 \pm .66$
7185	.0960	.1009	-5.1	$2.5 \pm .79$
8254	.0808	.0848	-5.0	$2.8 \pm .33$
8300	.0824	.0869	-5.5	$2.7 \pm .42$
TERTIARY ($>\text{CH}$)				
6950	0.0571	0.0618	-8.2	1.1 ± 0.86
8160	.0643	.0603	+6.2	$2.4 \pm .45$
AROMATIC ($>\text{CH}$)				
5900	0.535	0.520	-----	-----
8680	.0263	.0318	-----	-----
8750	.0097	.0169	-----	-----

The results are in excellent agreement for a compound having these values: $n_a=4$, $n_b=3$, $n_c=2$. The compound is therefore an isononane having two branches not joined to the same carbon atom. The first conclusion is reached from the number of methyl ($-\text{CH}_3$) groups, while the latter is concluded from the number of secondary ($>\text{CH}_2$) and tertiary ($>\text{CH}$) groups. The unknown must thereby be a dimethylheptane or a methylethylhexane where the two substitutions are on different carbon atoms.

Actually, the given isononane has been lately identified as 2,6-dimethylheptane, on the basis of a comparison of its boiling point, freezing point, density, and refractive index with those of a synthetic sample of 2,6-dimethylheptane prepared by Calingaert and Soroos in the laboratories of the Ethyl Gasoline Corporation.²⁰

3. LIMITATIONS OF THE METHOD

In connection with the application of the present method to unknown hydrocarbons, the following limitations should be pointed out: (a) Because of the assumed experimental uncertainty of about ± 8 percent, the uncertainty in the deduced value for the number of a given kind of group may be more than one if the number of the given kind of group is large or if the group is of low absorbing power. (b) The assumption of the equivalence of various isomeric structures known to differ somewhat in their absorption is a weakness of the method. However, unless a compound is an extreme example of such a type with the consequent multiplication of a minor effect, a correct analysis can still be made. (c) The possibility always exists that a compound of radically different structure not yet measured would reproduce the absorption curve of one of the substances included in or extrapolated from this study. Although the use of "check" points minimizes this possibility, the further accumulation of data on different types of hydrocarbons will materially increase the reliability of this method.

The development of the ideas presented in this paper was materially aided by the suggestions and guidance of O. R. Wulf, Bureau of Chemistry and Soils, United States Department of Agriculture, and of F. D. Rossini, Director of the American Petroleum Institute Research Project 6.

The author acknowledges the aid of F. W. Melpolder and P. Schoonover in many of the calculations.

VII. REFERENCES

- [1] E. W. Washburn. *Ind. Eng. Chem.* **25**, 891 (1933).
- [2] U. Liddel and C. Kasper. *BS J. Research* **11**, 599-618 (1933) RP610.
- [3] E. W. Washburn. *Ind. Eng. Chem.* **22**, 985-8 (1930).
- [4] G. H. B. Davis and E. N. McAllister. *Ind. Eng. Chem.* **22**, 1326 (1930).
- [5] L. A. Mikeska. *Ind. Eng. Chem.* **28**, 970-983 (1936).
- [6] G. Hugel. *Chimie & industrie* **26**, 1282 (1931).
- [7] E. W. Dean and G. H. B. Davis. *Chem. Met. Eng.* **36**, 618 (1929).
- [8] G. H. B. Davis, M. Lapeyrouse, and E. W. Dean. *Oil & Gas J.* **30**, 46, 92 (1932).

²⁰ The description of the synthetic compound and the identification of the isononane from petroleum will appear in a forthcoming paper in the *Journal of Research of the National Bureau of Standards*, by White and Rose, in collaboration with Calingaert and Soroos.

- [9] J. C. Vlughter, H. I. Waterman, and H. A. Van Westen. *J. Inst. Petroleum Tech.* **18**, 735 (1932); **21**, 661 (1935); **21**, 707 (1935).
- [10] J. C. Vlughter. Dissertation, Delft (1932).
- [11] H. I. Waterman. *Proc. World Petroleum Cong.* **2**, 332 (1933).
- [12] H. I. Waterman, P. C. Kruijff, W. J. K. Schönlaui, and A. J. Tulleners. *J. Inst. Petroleum Tech.* **20**, 159 (1934).
- [13] G. H. Fuchs and A. P. Anderson. *Ind. Eng. Chem.* **29**, 319 (1936).
- [14] A. L. Ward and W. H. Fulweiler. *Ind. Eng. Chem., Anal. Ed.* **6**, 396 (1934).
- [15] J. B. Hill and H. B. Coates. *Ind. Eng. Chem.* **20**, 641 (1928).
- [16] S. S. Kurtz, Jr., and A. Q. Ward. *J. Franklin Inst.* **222**, 563 (1936).
- [17] F. S. Brackett. *Proc. Nat. Acad. Sci.* **14**, 857-864 (1928).
- [18] F. W. Rose, Jr. *J. Research NBS* **19**, 143-161 (1937) RP1017.
- [19] W. Herschel. *Phil. Trans.* **90**, 225, 293, 437 (1800).
- [20] W. Abney and B. Festing. *Phil. Trans.* **172**, 887 (1881).
- [21] H. Julius. *Verhandl. Akad. Wetenschappen Amsterdam* **1**, 1 (1892).
- [22] W. W. Coblenz. *Investigations of Infrared Spectra*, Carnegie Inst., Wash. Pub. 35 (1905).
- [23] B. Donath. *Ann. Physik* [3] **58**, 609 (1896).
- [24] Ransohoff. Inaug. Dissertation, Berlin (1896).
- [25] L. Puccianti. *Nuovo Cimento* **11**, 241 (1900); *Physik. Z.* **1**, 49, 494 (1899-1900).
- [26] M. J. Lecomte. *Structure des molécules et spectres d'absorption dans l'infrarouge*, *Traité de Chimie Organique*, **III** (Masson et Cie., Paris, 1936).
- [27] R. B. Barnes. *Rev. Sci. Inst.* **7**, 265 (1936).
- [28] R. B. Barnes, L. G. Bonner, and E. U. Condon. *J. Chem. Phys.* **4**, 772 (1936).
- [29] M. P. Lambert and M. J. Lecomte. *Ann. Comb. Liquides*, 1001-1084 (1931); 979-993 (1934); 1077-1092 (1935).
- [30] V. Henri. *Etudes de Photochimie* (Gauthier-Villars et Cie, Paris, 1919).
- [31] G. B. Bonino. *Trans. Faraday Soc.* **25**, 876 (1929).
- [32] F. K. Bell. *J. Am. Chem. Soc.* **50**, 2940 (1928).
- [33] M. R. Freymann. *Compt. rend.* **193**, 657 (1931).
- [34] U. Liddel and O. R. Wulf. *J. Am. Chem. Soc.* **55**, 3574 (1933).
- [35] F. D. Rossini. *J. Research NBS* **13**, 21 (1934) RP692; **13**, 189 (1934) RP701.
- [36] O. R. Wulf and U. Liddel. *J. Am. Chem. Soc.* **57**, 1464 (1935).
- [37] O. R. Wulf. Paper 28 (Division of Physical and Inorganic Chemistry), Rochester, N. Y., meeting of the American Chemical Society, Sept. 6, 1937.
- [38] D. Williams and R. Taschek. *J. Applied Phys.* **8**, 497 (1937).
- [39] J. Lipka. *Graphical and Mechanical Computation* (John Wiley and Sons, Inc., New York, N. Y., 1918).
- [40] R. B. Barnes. *Phys. Rev.* **35**, 1524 (1930); **36**, 296 (1930).
- [41] O. R. Wulf, U. Liddel, and S. B. Hendricks. *J. Am. Chem. Soc.* **58**, 2287 (1936).
- [42] J. D. White and F. W. Rose, Jr. *J. Research NBS* **17**, 943 (1936) RP955.
- [43] L. Pauling. *J. Am. Chem. Soc.* **58**, 94 (1936).
- [44] B. M. Bloch and J. Errera. *J. phys. radium* **6**, 154 (1935).
- [45] R. Mecke, P. Gänswein, O. Vierling. *Z. Electrochem.* **40**, 474 (1934).
- [46] R. Mecke. *Angew. Chem.* **48**, 320 (1935).
- [47] E. Bartholomé and E. Teller. *Z. physik. Chem. [B]* **19**, 366 (1932).
- [48] R. Davis and K. S. Gibson. *NBS Misc. Pub. M114* (1931).

WASHINGTON, September 28, 1937.